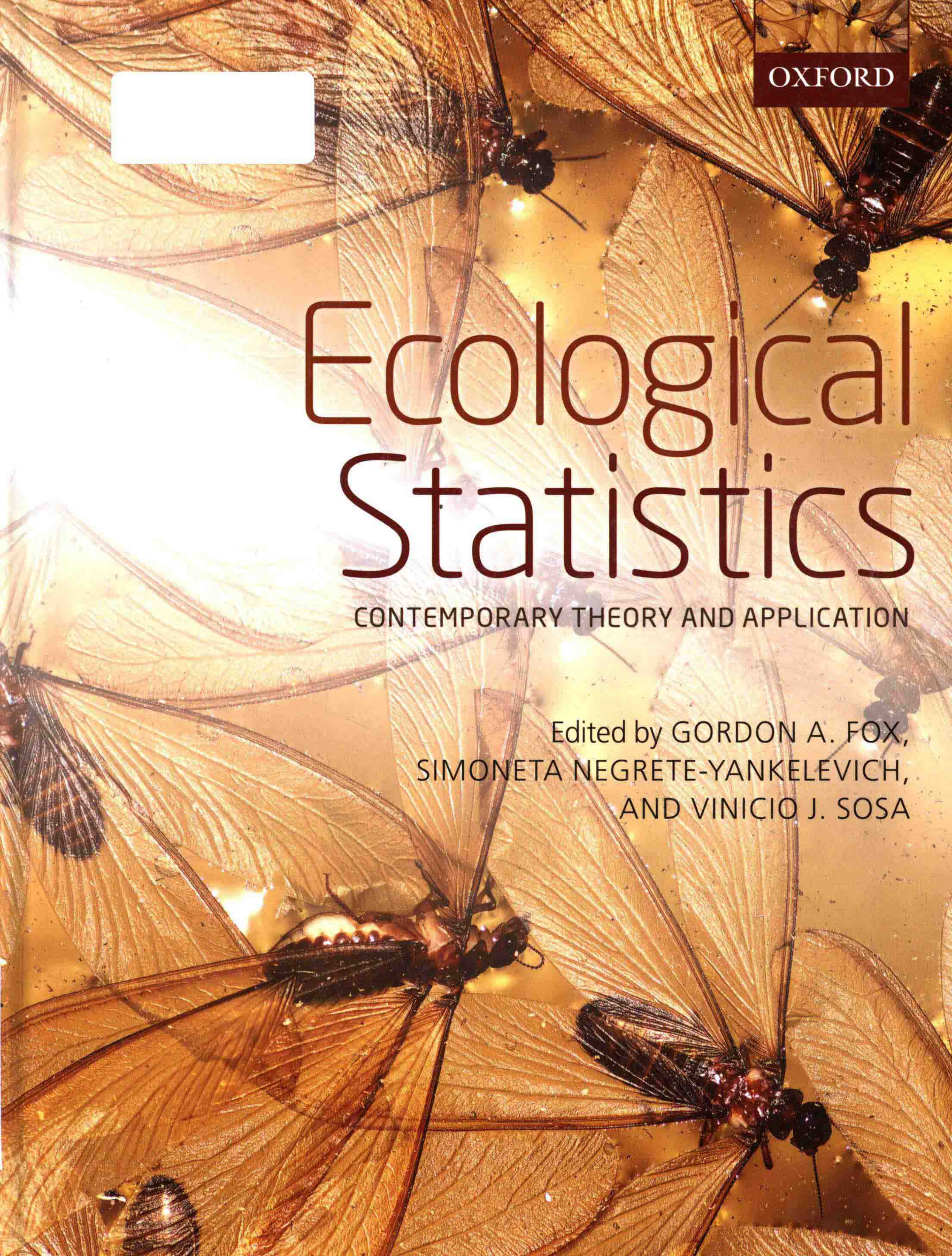

OXFORD

Ecological Statistics

CONTEMPORARY THEORY AND APPLICATION

Edited by GORDON A. FOX,
SIMONETA NEGRETE-YANKELEVICH,
AND VINICIO J. SOSA

Ecological Statistics: Contemporary Theory and Application

Edited By

GORDON A. FOX

University of South Florida

SIMONETA NEGRETE-YANKELEVICH

Instituto de Ecología A. C.

VINICIO J. SOSA

Instituto de Ecología A. C.

OXFORD
UNIVERSITY PRESS

Ecological Statistics: Contemporary Theory and Application. First Edition. Edited by Gordon A. Fox, Simoneta Negrete-Yankelevich, and Vinicio J. Sosa.
Published in 2015 by Oxford University Press.

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2014956959

ISBN 978–0–19–967254–7 (hbk.)
ISBN 978–0–19–967255–4 (pbk.)

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Ecological Statistics: Contemporary Theory and Application

Dedication

Gordon A. Fox – To Kathy, as always.

Simoneta Negrete-Yankelevich – A Laila y Aurelio, con amor infinito.

Vinicio J. Sosa – To Gaby, Eras and Meli.

Acknowledgments

The contributors did much more than write their chapters; they provided invaluable help in critiquing other chapters and in helping to think out many questions about the book as a whole. We would like to especially thank Ben Bolker for his thinking on many of these questions. Graciela Sánchez Ríos provided much-needed help with the bibliography. Fox was supported by grant number DEB-1120330 from the U.S. National Science Foundation. The Instituto de Ecología A.C. (INECOL) encouraged this project from beginning to end, and gracefully allocated needed funding (through the *Programa de Fomento a las Publicaciones de Alto Impacto/Avances Conceptuales y Patentes* 2012) to allow several crucial work meetings of the editors; without this help this book would probably not have seen the light of day.

List of contributors

Benjamin M. Bolker
Departments of Mathematics & Statistics and Biology
McMaster University
1280 Main Street West
Hamilton, Ontario L8S 4K1
Canada
bolker@mcmaster.ca

Yvonne M. Buckley
School of Natural Sciences
Trinity College , University of Dublin
Dublin 2
Ireland
buckleyy@tcd.ie
and
The University of Queensland
School of Biological Sciences
Queensland 4072
Australia

Gordon A. Fox
Department of Integrative Biology (SCA 110)
University of South Florida
4202 E. Fowler Ave.
Tampa, FL 33620
USA
gfox@usf.edu

James B. Grace
US Geological Survey
700 Cajundome Blvd.
Lafayette, LA 70506
USA
gracej@usgs.gov

Jessica Gurevitch
Department of Ecology and Evolution
Stony Brook University
Stony Brook, NY 11794-5245
USA
Jessica.Gurevitch@stonybrook.edu

Bruce E. Kendall
Bren School of Environmental Science & Management
University of California, Santa Barbara
Santa Barbara CA 93106-5131
USA
kendall@bren.ucsb.edu

Marc J. Lajeunesse
Department of Integrative Biology (SCA 110)
University of South Florida
4202 E. Fowler Ave.
Tampa, FL 33620
USA
lajeunesse@usf.edu

Michael A. McCarthy
School of BioSciences
The University of Melbourne
Parkville VIC 3010
Australia
mamcca@unimelb.edu.au

Earl D. McCoy
Department of Integrative Biology (SCA 110)
University of South Florida
4202 E. Fowler Ave.
Tampa, FL 33620
USA
edm@mail.usf.edu

Shinichi Nakagawa
Department of Zoology
University of Otago
340 Great King Street
P.O. Box 56
Dunedin
New Zealand
shinichi.nakagawa@otago.ac.nz
and
School of Biological, Earth and Environmental Sciences
University of New South Wales
Sydney
NSW 2052
Australia

Simoneta Negrete-Yankelevich
Instituto de Ecología A. C. (INECOL)
Carretera Antigua a Coatepec 351
El Haya Xalapa 91070

Veracruz
México
simoneta.negrete@inecol.mx

Jonathan R. Rhodes
The University of Queensland
School of Geography, Planning, and Environmental Management
Brisbane
Queensland 4072
Australia
jrhodes@uq.edu.au

Shane A. Richards
School of Biological & Biomedical Sciences
Durham University
South Road
Durham, DH1 3LE
UK
s.a.richards@durham.ac.uk

Samuel M. Scheiner
Division of Environmental Biology
National Science Foundation
Arlington, VA 22230
USA
sscheine@nsf.gov

Donald R. Schoolmaster Jr.
U. S. Geological Survey
700 Cajundome Blvd.
Lafayette, LA 70506
USA
schoolmasterd@usgs.gov

Vinicio J. Sosa
Instituto de Ecología A. C. (INECOL)
Carretera Antigua a Coatepec 351
El Haya Xalapa 91070
Veracruz
México
vinicio.sosa@inecol.mx

Contents

Introduction

Vinicio J. Sosa, Simoneta Negrete-Yankelevich, and Gordon A. Fox

Why another book on statistics for ecologists?

This is a fair question, given the number of available volumes on the subject. The reason is deceptively simple: our use and understanding of statistics has changed substantially over the last decade or so. Many contemporary papers in major ecological journals use statistical techniques that were little known (or not yet invented) a decade or two ago. This book aims at synthesizing a number of the major changes in our understanding and practice of ecological statistics.

There are several reasons for this change in statistical practice. The most obvious cause is the continued growth of computing power and the availability of software that can make use of that power (including, but by no means restricted to, the R language). Certainly, the notebook and desktop computers of today are vastly more powerful than the mainframe computers that many ecologists (still alive and working today) once had to use. Both hardware and software can still impose limits on the questions we ask, but the constraints are less severe than in the past.

The ability to ask new questions, together with a growing body of practical experience and a growing cadre of ecological statisticians, has led to an increased level of statistical sophistication among ecologists. Today, many ecologists recognize that the questions we ask should be dictated by the scientific questions we would like to address, and not by the limitations of our statistical toolkit. You may be surprised to hear that this has ever been an issue, but letting our statistical toolkit determine the questions we address was a dominant practice in the past and is still quite common. However, increasingly today we see ecologists adapting procedures from other disciplines, or developing their own, to answer the questions that arise from their research. This change in statistical practice is what we mean by "deceptively simple" in the first paragraph: the difference between ecologists' statistical practice today and a decade or two ago is not just that we can compute quantities more quickly, or crunch more (complex) data. We are using our data to consider problems that are more complex. For example, a growing number of studies use statistical methods to estimate parameters (say, the probability that the seed of an invasive pest will disperse X meters) for use in models that consider questions like rates of population growth or spread, risks of extinction, or changes to species' ranges; fundamental questions, but ones that were previously divorced from statistics. Meaningful estimates of these quantities require careful choice of statistical approaches, and sometimes these approaches cannot be

Ecological Statistics: Contemporary Theory and Application. First Edition. Edited by Gordon A. Fox, Simoneta Negrete-Yankelevich, and Vinicio J. Sosa.
Published in 2015 by Oxford University Press.

limited to the contents of traditional statistics courses. This is of course only a point in a continuum; future techniques will continue to extend our repertoire of tractable questions and new books like this will continue to appear.

There is nothing wrong with using basic or old statistical techniques. Techniques like linear regression and analysis of variance (ANOVA) are powerful, and we continue to use them. But using techniques because we know them (rather than because they are appropriate) amounts to fitting things into a Procrustean bed–it does not necessarily ask the question we want to ask. We encountered recently a small but illustrative example in one of our labs: identifying environmental characteristics predicting presence of a lily, *Lilium catesbaei* (Sommers et al. 2011). It seemed reasonable to approach this problem with logistic regression (GLM with a binomial link; chapter 6), using site characteristics as the predictors and probability of presence/absence as the outcome. In reviewing literature on prediction of site occupancy, we found that a very large fraction of studies used a very different approach: ANOVA to compare the mean site characteristics of occupied with unoccupied sites. These might seem like comparable approaches, but they are quite different: logistic regression models probability of occupancy as a function of site characteristics, while ANOVA considers occupancy to be like an experimental treatment that somehow causes site characteristics! Yet many studies had used just this approach. To explore the problem, we analyzed the data using both approaches. The set of explanatory variables that we found predicted lily presence (using logistic regression) was not the same as the set of predictors for which occupied and unoccupied sites differed significantly (using ANOVA). The difference is not because the two approaches differ in power, or because we strongly violated underlying assumptions using one of the methods; the different results occur because the questions asked by the two approaches are quite different. This underlines a point that is often not obvious to beginners: the same data processed with different methods leads to different answers. By choosing a statistical method because it is convenient, we run the risk of answering questions we do not intend to ask. Worse still, we may not even realize that we have answered the wrong question.

The idea for this book emerged during a couple of occasions on which Fox came to Mexico to teach a survival module in the Sosa–Negrete statistics course for ecology graduate students. Dinner conversations often converged on the conclusion that, despite considerable efforts, learning statistics continues to be boring for many ecologists and more often than not, it feels a bit like having dental work done: frightening and painful but necessary for survival.

However, nothing could be further from the truth. Statistics is at the core of our science, because it provides us with tools that help us interpret our complex (and noisy) picture of the natural world (figure I.1). Ecologists today are leading in the development of a number of areas of statistics, and potentially we have a lot more to contribute. Many techniques used by ecologists are thoughtful, efficient, powerful, and diverse. For young ecologists to be able to keep up with this phenomenal advance, old ways of teaching statistics (based on memorizing which ready-made test to use for each data type) no longer suffice; ecologists today need to learn concepts enabling them to understand overarching themes. This is especially clear in the contribution that ecologists and ecological problems have made to the development of roll-your-own models (Hilborn and Mangel, 1997; Bolker, 2008).

The chapters of this book are by experienced ecologists who are actively working to upgrade ecologists' statistical toolkit. This upgrade involves developing models and statistical techniques, as well as testing the utility, usability, and power of these techniques in real ecological problems. Some of the techniques highlighted in the book are not new, but are underused in ecology, and can be a great aid in data analysis.